Temperature effects in metal-clad semiconductor nanolasers

Abstract: As the field of semiconductor nanolasers becomes mature in terms of both the miniaturization to the true sub-wavelength scale, and the realization of room temperature devices, the integrated treatment of multiple design aspects beyond pure electromagnetic consideration becomes necessary to further advance the field. In this review, we focus on one such design aspect: temperature effects in nanolasers. We summarize recent efforts in understanding the interplay of various temperature-dependent parameters, and study their effects on optical mode and emission characteristics. Building on this knowledge, nanolasers with improved thermal performance can be designed, and their performance evaluated. Although this review focuses on metal-clad semiconductor lasers because of their suitability for dense chip-scale integration, these thermal considerations also apply to the broader field of nanolasers.

Keywords: Design co-optimization; nanoscale devices; semiconductor lasers; temperature effects.

DOI 10.1515/nanoph-2013-0058
Received December 16, 2013; accepted September 12, 2014

1 Introduction

In the past decade, lasing has been demonstrated in numerous wavelength and sub-wavelength scale structures, including dielectric micro-discs [1–4], photonic crystals [5–9], nanowires [10, 11] and nanorods [12], nano-membranes [13–15], micro-pillars [16–18], and metal-clad nano-cavities [19–24]. While all these devices enable fundamental research of various nanoscale phenomena [25–27], the design and analysis of nanolasers have focused almost exclusively on the optical mode, i.e., pure electromagnetic (EM) consideration, usually at 4.5 K, 77 K and room temperature. The experimental demonstrations have therefore focused on validating the optical cavity design and showing lasing behavior. As the field of nanolasers becomes more mature, and continues developing toward stable devices suitable for on-chip integration, other interdependent aspects of nanolaser design will need to be considered. In this review, we summarize recent efforts toward the understanding and management of one such design aspect, nanolaser temperature dependence. While thermal dynamics in vertical-cavity surface-emitting lasers (VCSELs) has been studied in depth [28, 29], it has been largely overlooked in nanolasers. However, thermal effects can have a profound impact on a nanolaser’s performance, both as a potential failure mechanism and through the temperature dependence of material parameters such as gain spectrum and cavity mode behavior. In studying nanolaser thermal behavior, we seek to accomplish two primary goals, namely, to understand the interplay of temperature-dependent effects on nanolasers, and to use this knowledge to design and evaluate nanolasers with improved thermal performance.

To understand the interconnected effects of temperature on nanolasers, it is first important to know the temperature dependence of the nanolaser’s material parameters. Through their effect on the nanolaser’s gain spectrum, material loss, and cavity resonance wavelength, the temperature-dependent material parameters also have an effect on the nanolaser’s quantum electrodynamics (QED) effects. Two interconnected QED parameters, both temperature-dependent, become increasingly prominent as the device size is reduced to the nanoscale. They are 1) the Purcell factor, $F_p$, which characterizes the enhancement or inhibition of the spontaneous emission rate in a cavity compared to that in free space, and 2) the
spontaneous emission factor, $\beta$, which is defined as the ratio of spontaneous emission channeled into the lasing mode versus all other modes. Both parameters are considered measures of nanolaser efficiency, especially for below-threshold carrier-photon dynamics. Their effects have proven crucial to recent work including diode lasers with greater modulation bandwidth [30–32], energy efficiency [7, 33], and absence of a threshold [24, 34]. A recent study [35] of the temperature dependence of $F_\beta$, and in turn, $\beta$, revealed that the effects of temperature can be dramatic: for example, a laser designed for high $\beta$ at room-temperature may exhibit low $\beta$ at low temperatures and vice versa. Making a measurement of $\beta$ at low temperatures and/or designing a high-$\beta$ laser therefore requires a predictive model that may account for these changes.

With knowledge of these nanolaser temperature dependencies, it becomes possible to evaluate nanolaser designs for their temperature performance, as well as to make modifications to improve thermal management, and evaluate the effects of these modifications. An analysis of a device’s temperature performance requires calculation of its self-heating, such as the simulation performed by Liu et al. [36] for a continuous wave (CW) optically-pumped device, or by Gu et al. [23] for a CW electrically-pumped device. This self-heating, tempered by the nanolaser’s ability to dissipate heat, can cause a significant temperature rise in the semiconductor, which in turn affects material parameters and therefore the nanolaser’s QED effects as well as optical behavior. Analysis of the nanolaser’s performance, therefore, should take place at the actual internal operating temperature, rather than at the ambient temperature. With the ability to do this interdependent analysis, it is then possible to evaluate the effects of design modifications that reduce the amount of heat generated, or increase the nanolaser’s ability to dissipate heat.

Here, we review recent progress toward understanding the effects of temperature on nanolaser performance, and toward using this understanding to improve device thermal management. While the temperature-dependent considerations outlined above are common to all types of nanoscale devices, we focus here on the example of metal-clad semiconductor nanolasers. Among the numerous nanolaser configurations, metal-clad nanolasers make particularly strong candidates for densely packed arrays of individually addressable coherent sources for on-chip applications, thanks to their compact size both in device footprint and mode confinement, as well as the lack of optical interference. For this reason, since the demonstration of the first metallic-coated nanolaser, which operated at 77 K under pulsed electrical injection [19], the metal-clad nanolaser has attracted a significant amount of attention. We discuss research efforts on temperature effects in metal-clad semiconductor nanolasers, with the applications of temperature varying operations and dense chip-scale integration in mind.

The optical modes of metal-clad cavities can be grouped into two main categories, namely surface plasmon polariton (SPP) modes and photonic modes. Although highly confined, SPP modes have the disadvantage of high loss at telecommunication wavelengths, as a result of the relatively large mode overlap of the optical field with the metal. Nanolasers utilizing SPP modes have been demonstrated at cryogenic temperatures, where sufficient gain can be attained [19, 37]. However, the high threshold gain of SPP modes in such cavities has made room temperature operation challenging. To this end, various attempts have been made to mitigate both the high absorptive metal loss and radiation loss. For example, improved cavity feedback mechanism that reduces the radiation loss, combined with hybridization of SPP and photonic modes which mitigates the metal loss were employed to achieve room temperature operation in optically pumped surface plasmon nanolasers [38].

In the case of photonic modes, the mode overlap with the metal is usually much smaller than that of SPP modes. Hence, photonic lasing modes typically have higher Q-factors and lower threshold gain values, albeit at the expense of reduced mode confinement. For metal-clad nanolasers based on photonic modes, metal losses may be minimized using a metallo-dielectric composite structure, in which a shield layer is placed between the active region and the metal cladding [39]. From a purely electromagnetic perspective, the optimal shield thickness for a given total device size, corresponding to minimal threshold gain of a photonic mode, can be determined numerically [39], and may be approximated analytically [40]. Lasing in an optically-pumped metallo-dielectric subwavelength laser, with such an optimized thickness of SiO$_2$ shield, was demonstrated at room temperature [41]. In the case of electrical injection, the dielectric also serves as the electrical insulation layer and the passivation layer. CW operation in this type of nanolaser was later demonstrated at 140 K under electrical pumping [21], and, most recently, at room temperature [22]. We focus our discussion here on metallo-dielectric nanolasers, as a promising avenue toward dense chip-scale integration at room temperature and higher. Although the metal cladding and the rotational symmetric emission from the subwavelength aperture of these nanolasers pose difficulties in coupling the emitted light into integrated waveguides, it has been proposed that efficient out-coupling into an integrated Si waveguide.
2 Material gain and electromagnetic cavities for nanolasers: temperature effects

2.1 Thermal dependence of material gain

To understand experimental nanolaser performance at various temperatures, a temperature dependent model for the material gain and EM cavity modes is essential. For the lasers discussed in this review, the active materials are multi-quantum wells (MQW) of 1.6Q/1.3Q InGaAsP (In0.56Ga0.43As0.933P0.067/In0.737Ga0.263As0.569P0.431) for the optically pumped devices, and bulk In0.53Ga0.47As for the electrically pumped devices. The gain spectra of these materials, calculated according to the semiclassical model of [44], are shown in Figure 1, for various temperatures, at carrier densities of 2.0×10¹⁸ cm⁻³ and 7.07×10¹⁸ cm⁻³, respectively. The reason for plotting the material gain with these parameters will become clear in the following sections. For now, we focus on the main trends. Generally, for a fixed carrier density, the material gain decreases with rising temperature. This may be explained if we consider the total carrier density \( N \) to be the sum of thermally excited carriers \( N_{\text{therm}} \) and externally injected carriers (either through optical or electronic injection) \( N_{\text{ext}} \). As the temperature rises, \( N_{\text{therm}} \) increases, which for a fixed \( N \) means that \( N_{\text{ext}} \) decreases, and hence the inversion factor, \( f_{\text{inv}} \), decreases. The material gain is directly proportional to \( f_{\text{inv}} \), resulting in a lower peak gain at higher temperatures. Additionally, for a fixed \( N \), the gain spectra for both material systems red shift as the temperature rises. This is primarily due to the fact that the bandgap energy decreases with increasing temperature [45]. Finally, for a fixed \( T \), as the carrier density increases the spectra broaden and the peak gain blue-shifts [44].

2.2 Thermal dependence of cavity resonance

The resonant wavelengths of the cavity modes of metal-clad nanolasers also depend upon temperature, although to a much lesser extent than the gain spectrum. The thermo-optic coefficient (TOC), \( d n_r / d T \), in InGaAs and InGaAsP is on the order of \( 10^{-4} \) RIU/K [46], where RIU refers to refractive index units. This means that a change in temperature by 100 K only changes the refractive index by a factor of 0.01. In Figure 2, we plot both the real and imaginary part of silver’s permittivity at 1550 nm via a temperature dependent Drude model, scaled to match

![Figure 1](image-url)  
*Figure 1* Material gain spectra with temperature as a parameter for (A) 10 nm 1.6Q/1.3Q InGaAsP QW near the transparent carrier density and (B) bulk In0.53Ga0.47As near the threshold carrier density.
the empirical data of Johnson and Christy [47] at room temperature [35]. Similarly to the semiconductor behavior, the real part of the permittivity of the metal cladding varies little with respect to temperature. On the other hand, the imaginary part of the permittivity of the metal cladding may increase by an order of magnitude with a 100 K increase in temperature; the most dramatic part of this increase occurs from around 100 K – 130 K. As it turns out, even this large change in imaginary permittivity has little effect on the resonant wavelength, since the effective index of a waveguide mode is essentially invariant with respect to the cladding permittivity [40]. Temperature, therefore, should have the greatest effect on the metal absorption loss experienced by the cavity modes, and a modest effect on resonant wavelength.

A rigorous calculation using finite-element modeling, empirical room-temperature optical properties, and the temperature-dependent Drude model verifies that the variation of cavity resonant wavelength is small over the temperature range 75 K – 400 K [35]. As an example, we simulate a nanolaser geometry similar to that shown in Figure 3A, but with a gain radius of 250 nm and optimal shield thickness of 100 nm, using the commercial software COMSOL Multiphysics’ EM module. In Figure 3D, the resonant wavelengths of the first six cavity modes are plotted as a function of temperature, for both a positive [46] and effectively negative TOC [48, 49] of the gain medium. The latter corresponds to the empirical observations of Massum et al. [48] and Ramoo et al. [49]. In these works on InGaAsP-based cavities, the resonant wavelength was observed to increase with the temperature. However, the analysis did not consider the temperature dependence of the individual materials; hence our use of the term “effectively” negative TOC. The lasing mode of the 250 nm cavity is the TE \text{011} \text{ mode; for the case of positive (effectively negative) TOC, its resonant wavelength ranges...}
from 1.465 μm (1.568 μm) at 77 K–1.514 μm (1.468 μm) at 400 K.

2.3 Thermal effects on the optimum shield thickness

One aspect of metallo-dielectric nanolaser design is the choice of dielectric shield thickness. The optimum shield thickness for a given total device diameter can be calculated using the numerical technique outlined in [39], or by the analytical approximations outlined in [40]. Both techniques involve first finding the optimal shield thickness in a composite gain 2D waveguide, which shares the same optimal shield thickness with its 3D counterpart. Such a waveguide is described in the conceptual diagram in Figure 3B, where the upper and lower waveguide sections extend to infinity. The mode is only supported in the central region where the high-index gain material resides. Figure 3C shows the numerically determined optimal SiO$_2$ shield thickness and its corresponding threshold gain for this infinite waveguide, for a total device radius of 460 nm [39]. This optimal shield thickness depends on the real part of the material permittivities, which depend weakly on temperature, as discussed in Section 2.2. However, it would intuitively seem that the large increase in metal loss with increasing temperature would necessitate a large increase in optimum shield thickness, to isolate the mode from the increasingly lossy metal; this proves not to be the case. In fact, the optimum shield thickness is virtually temperature-independent [40]. However, this optimum shield thickness only takes electromagnetic considerations into account. The choice of shield thickness and material will impact the laser’s ability to dissipate heat, as we show in Section 4, and through its effect on the laser’s internal operating temperature, will dynamically affect the material gain spectrum, as well as metal loss.

3 Purcell effect evaluation and the temperature dependence of the spontaneous emission factor in semiconductor nanolasers

3.1 Purcell effect in semiconductor nanolasers

Apart from understanding the temperature dependence of material gain and electromagnetic cavities, the understanding of the temperature dependence of the Purcell factor $F_p$ and the spontaneous emission factor $\beta$ is equally important, largely motivated by the quest for energy-efficient operations in nano-scale devices. While the concept of thresholdless operation continues to be a subject of debate [25, 50], the modulation and efficiency improvements enabled by wavelength-scale cavities, which are directly related to the modification of the spontaneous emission (Purcell effect) in these cavities, is fairly well understood [30–32]. For example, with proper design, the cavity of a sub-wavelength laser may be designed such that most of the spontaneous emission is channeled into the lasing mode. In so doing, unwanted emission into non-lasing modes is mitigated, and the below-threshold efficiency is limited only by non-radiative recombination. In nano-scale lasers, enhanced emission together with a reduced number of cavity modes relative to large lasers can have significant effects, especially on sub-threshold behavior. These effects are generally desirable, as they tend to increase the utilization of spontaneous emission into the lasing mode, thus lowering the lasing threshold and increasing $\beta$. If the desired cavity mode has the highest Purcell factor amongst all cavity modes, a high $\beta$ laser can be realized even in a multi-mode cavity. With this design goal in mind, it is important to accurately evaluate the $F_p$ of all cavity modes, taking into account the emitter environment and the semiconductor gain material properties. While the original Purcell effect evaluation was for radio frequency micro-cavities, the formal treatment of Purcell effect specific to nano-scale devices wasn’t presented until the work of [26, 32, 51]. Tailored for nanolasers, these studies provide insight into the fundamental physics in the nano-cavities, although the temperature dependence of $F_p$ was not included.

For a cavity of quality factor $Q$, active region volume $V_a$ and refractive index $n$, emitting at the free-space wavelength $\lambda$, the Purcell factor $F_p$ describes the rate of spontaneous emission relative to the emission in bulk. In the simplified expression of a two-level system and a single cavity mode that are on resonance and spatially overlapping, $F_p$ usually takes the form

$$F_p = \frac{3}{4\pi^2 V_a} \left( \frac{\lambda}{n} \right)^3.$$  \hspace{1cm} (1)

This simplified form is the one most commonly seen and used in the literature [52] to quantify the cavity-enhanced or -inhibited rate of spontaneous emission, relative to the emission in freespace. However, using the emitter-field-reservoir model in the quantum theory of damping, effects ignored in the commonly used Purcell
factor expression can be captured. The result revealed that, this expression is far from accurate in semiconductor nanolasers with bulk or MQW gain media.

We apply the results from the non-relativistic QED treatment of 2-level systems to a 3-level laser, in which emitters are pumped from the ground state $|1\rangle$ to an excited state $|3\rangle$ and quickly decay from state $|3\rangle$ to a lower state $|2\rangle$; the lasing transition is between states $|2\rangle$ and $|1\rangle$ [26]. Semiconductor lasers in particular are frequently modeled in this manner, even though their underlying physics differs: state $|2\rangle$ describes the condition where a conduction band state is occupied and the valence band state of the same crystal momentum is vacant, while state $|1\rangle$ describes the condition when the conduction band state is vacant and the valence band state is occupied [53]. To describe such a system, we construct a basic model similar to that in ([54], §9) and [55]. We suppose each emitter to interact with all modes of the cavity, but ignore direct interaction among emitters. The cavity modes, on their part, undergo damping as a result of loss at the cavity boundaries, and we model the loss as a thermal reservoir.

Loss at the cavity boundary, such as loss in a metallic mirror, or loss of energy through the mirror and its eventual conversion to heat at some remote point in space, generally satisfies the assumptions of a reservoir model: it is weak interaction with a large stochastic system that is disordered and does not retain memory of past interactions. Further, this reservoir is passive, as it does not return energy to the mode. Rather, it drains the mode energy over time and is commonly known as the zero temperature condition. The Hamiltonian describing each single emitter in this system can be expressed as

$$\hat{H} = \hat{H}_e + \hat{H}_r + \hat{H}_{ae} + \hat{H}_{fe} + \hat{H}_{FR}$$

where $\hat{H}_e$, $\hat{H}_r$, and $\hat{H}_{ae}$ are the emitter, field and reservoir Hamiltonian, respectively. $\hat{H}_{fe}$ denotes interaction between the emitter and the field modes, while $\hat{H}_{FR}$ denotes interaction between the field modes and the reservoir.

We note that even if, by assumption, a given emitter does not directly interact with other emitters, the field modes still interact with all emitters present, rather than only with a single emitter. This interaction is not included in the Hamiltonian in Eq. (2), either explicitly or as part of the reservoir. In a solid-state system where an emitter interacts with the field, and the field interacts with a thermal reservoir, the cavity Purcell factor $F_p$ is defined as the ratio of spontaneous emission in a cavity to that in bulk material of effective index $n_e$ with no cavity [56, 57].

The spontaneous emission probability in the bulk material, $P_{\text{material}}^{\text{bulk}}(2\rightarrow1, \omega_0)$, takes the same form as in free space, except that $e_0$ is replaced by the permittivity of the medium $\varepsilon_r = n_e^2 e_0$ and that $c$ is scaled down by the refractive index $n_e$. It is expressed as

$$P_{\text{material}}^{\text{bulk}} = \frac{\omega_0^3}{3\pi \varepsilon_0 (c / n_e)^3} \tau_{\text{coll}} |\mathbf{p}(\omega_0)|^2 D(\omega_0) d\omega_0$$

$$= \frac{1}{3\pi |\mathbf{r}|^2 \tau_{\text{coll}} |\mathbf{p}(\omega_0)|^2}$$

where $\omega_0$ is the mode resonant frequency, $|\mathbf{p}(\omega_0)|$ is the dipole matrix element, and $D(\omega_0)$ characterizes the inhomogeneity of the system. The intraband collision time, $\tau_{\text{coll}}$, is the average time between carrier-carrier and carrier-phonon collisions, and decreases with increasing temperature [58]. In the second line of Eq. (3), we evaluate $\omega_0$, and $|\mathbf{p}(\omega_0)|$ at the center frequency $\omega_{21}$ of the homogeneous broadening spectrum $D(\omega_0)$ and pull them out of the integration, because these quantities vary relatively little over the homogeneous broadening range.

In a damped cavity, the mode interacts with the reservoir. Provided that equilibrium between the mode and the reservoir is reached, we obtain the photonemission probability in steady-state,

$$P_{2\rightarrow1, \text{equilibrium}} = \sum_{i} \frac{\omega_i}{\hbar} (n_i + 1) \left[ \frac{1}{2} |\mathbf{p}(\omega_i)|^2 D(\omega_i) \right]$$

where $R(\omega-\omega_{21}, \tau_{\text{coll}})$ is the homogeneous broadening function and depends on $r_{\text{coll}}$. Viewed as a function of $\omega$, $R(\omega)$ peaks at $\omega_{21}$ has a width on the order of $1/\tau_{\text{coll}}$, and satisfies $\int R(\omega-\omega_{21}, \tau_{\text{coll}}) d\omega = 2\pi \tau_{\text{coll}}$ [59]. The Lorentzian $L_\omega(\omega-\omega_{21})$ in Eq. (4) is expressed as

$$L_\omega(\omega-\omega_{21}) = \frac{2Q}{\pi \omega_{21}^2} = \frac{2Q}{\pi \omega_{21}^2} \frac{\left( \frac{1}{2} \Delta \omega_k \right)^2}{\Delta \omega_k^2 + \left(\omega-\omega_{21}\right)^2}$$

and the quality factor is defined as $Q = \omega_1 / \Delta \omega_k$. From the discussion in Section 2, the quality factor is a strong function of temperature due to the dependence of the full-width-at-half-maximum (FWHM), $\Delta \omega_1$, on the imaginary part of the metal-cladding permittivity. This is especially
true for well-confined modes where the $Q$ is dominated by absorptive, rather than radiation, losses. The convolution in Eq. (4) determines the emission probability in a cavity for an inhomogeneously broadened ensemble of emitters, when the mode-reservoir equilibrium has been reached. The effect of the reservoir on the emission probability is described by $L_k(\omega - \omega_k)$, whose spectral property is described by Eq. (5).

The Purcell factor $F_p$ is then

$$F_p = \frac{P_{\text{em, equilibrium}}^{\text{material}}}{P_{\text{em, initial}}^{\text{material}}} = \sum_k \frac{3\pi e^2 (c/\hbar)^2}{\tau_{\text{coll}}} \left| \mathbf{\rho}_{kk} (\omega_k) \right|^2 \frac{1}{\rho_{kk} (\omega_k)} \frac{1}{V_s} \left| \mathbf{e}_k (\mathbf{r}) \right|^2 \int d^3 \mathbf{r} \quad \text{(6)}$$

The emission probability in Eq. (4), and hence the Purcell factor in Eq. (6), depends on the location $\mathbf{r}$ of the emitter. More precisely, it depends on the normalized mode field at the location of the emitter $\mathbf{e}_k (\mathbf{r})$, as well as on the orientation of the emitter's dipole moment matrix element $\mathbf{\rho}_{kk} (\omega_k)$ relative to the field. If the emitters are randomly oriented and uniformly distributed over an active region of volume $V_s$, the quantity $\left| \mathbf{\rho}_{kk} (\omega_k) \mathbf{e}_k (\mathbf{r}) \right|^2$ is replaced by its average over all locations and orientations.

$$\left| \mathbf{\rho}_{kk} (\omega_k) \mathbf{e}_k (\mathbf{r}) \right|^2 \rightarrow \frac{1}{3} \left| \mathbf{\rho}_{kk} (\omega_k) \right|^2 \frac{1}{V_s} \left| \mathbf{e}_k (\mathbf{r}) \right|^2 \int d^3 \mathbf{r} \quad \text{(7)}$$

where the coefficient 1/3 accounts for the random emitter orientation.

In certain situations, the carrier distribution over $V_a$ may become non-uniform. For example, in MQW structures, the carrier distributions in the well and barrier regions differ significantly. Even in bulk semiconductors, the recombination of carriers may vary spatially, with the highest rates occurring at field antinodes. This is the case if the recombination at field antinodes is so rapid that diffusion of carriers from other parts of the active volume is not fast enough to avoid depletion. Carrier depletion at field antinodes and subsequent diffusion from the nodes toward the antinodes leads to the spatial inhomogeneity of the recombination. At room temperature, the diffusion length of carriers in InGaAsP (i.e., average distance traveled before recombination) is on the order of 1–2 μm [60]. The distance between the field node and antinode in visible and near infra-red sub-wavelength semiconductor cavities, on the other hand, is usually <0.5 μm [19, 20]. Thus, the depletion regions would remain relatively depleted due to the finite diffusion time. Under these circumstances, Eq. (7) should then be replaced by an appropriately weighted average.

$$F_p = \sum_k \frac{\pi(c/\hbar)^2 \omega_k}{\tau_{\text{coll}}} \left| \mathbf{\rho}_{kk} (\omega_k) \right|^2 \frac{1}{V_s} \left| \mathbf{e}_k (\mathbf{r}) \right|^2 \int d^3 \mathbf{r} \quad \text{(8)}$$

where $\Gamma_k$ is the energy confinement factor of mode $k$. Equation (8) permits several observations. Firstly, the double integral in Eq. (8) is the convolution of inhomogeneous broadening $D(\omega_k)$, cavity Lorentzian $L_k(\omega - \omega_k)$, and homogeneous broadening $R(\omega - \omega_{\text{coll}}, \tau_{\text{coll}})$. It should be noted that although the homogeneous broadening function $R(\omega)$ and the inhomogeneous broadening function $D(\omega)$ appear symmetrically in Eq. (8), they may in principle exhibit different dynamics. In particular, rapid recombination of carriers near the mode frequency $\omega_k$ may deplete the carrier population at that frequency faster than it is replenished by intraband scattering (this phenomenon is known as “spectral hole burning”). In such cases, it could be meaningful to disaggregate the integral in $d\omega_{\text{coll}}$ in Eq. (8) and define separate Purcell factors for carriers at different frequencies $\omega_{\text{coll}}$ [61]. More typically, however, especially at room temperatures, the intraband relaxation time $\tau_{\text{coll}} \sim 0.3$ ps of InGaAsP is much shorter than photonemission time, and the distribution of carriers $D(\omega_{\text{coll}})$ is at all times the equilibrium distribution ([44], Appendix 6). This equilibrium distribution closely resembles the photoluminescence spectrum [62]. In semiconductor lasers utilizing bulk or MQW gain material, it is the broadest of the three convolution factors in Eq. (8) and therefore dominates the convolution. For InGaAsP at room temperature, the FWHM of $D(\omega_{\text{coll}})$ and $R(\omega - \omega_{\text{coll}}, \tau_{\text{coll}})$ are approximately $7 \times 10^{10}$ rad/s and $6.7 \times 10^{11}$ rad/s, respectively. $D(\omega_{\text{coll}})$ dominates the convolution in Eq. (8) as long as the cavity $Q$ factor is above 19, which corresponds to a FWHM of $7 \times 10^{13}$ rad/s. For practical cavities, the $Q$ factor will be significantly larger; thus diminishing the contribution of $L_k(\omega - \omega_k)$ to the resulting Purcell factor. In fact, $R(\omega - \omega_{\text{coll}}, \tau_{\text{coll}})$, alone, dominates $L_k(\omega - \omega_k)$ if the $Q$ factor is greater than 200 [58, 63]. Consequently, in typical III-V semiconductor lasers with MQW or bulk gain material, the cavity $Q$ factor plays a negligible role in determining the spontaneous emission rate and $F_p$. Further, while the cavity lineshape broadens with temperature for well-confined cavity modes, the homogeneous lineshape broadens as well. Secondly, $F_p$ may be large in small laser cavities due to its inverse proportionality to the active region volume $V_a$. However, $F_p$ is actually inversely proportional to the effective size of the mode, $V_a/\Gamma_k$, where the mode-gain overlap factor $\Gamma_k$ is defined in Eq. (8) and describes the spatial overlap between the mode and
the active region. Thus, if the mode is poorly confined, \( \Gamma_k < < 1 \), \( F_p \) will remain small, despite a small active region.

### 3.2 Purcell factor of metal-clad nanolasers operating at room-temperature

Using the optically pumped room temperature metallo-dielectric nanolaser reported in Nezhad et al. [20], whose geometry is similar to that in Figure 3, Figure 4 shows its lasing mode’s electric field profile and the three spectra in the evaluation of the Purcell factor: cavity lineshape \( L_k(\omega - \omega_k) \) (Figure 4B), homogeneous broadening \( R(\omega - \omega_{21}, \tau_{coll}) \) (Figure 4C), and inhomogeneous broadening \( D(\omega_{21}) \) approximated by the photoluminescence (PL) spectrum (Figure 4D). It is then evident that the role of \( L_k(\omega - \omega_k) \) and \( R(\omega - \omega_{21}, \tau_{coll}) \) are diminished in the determination of Purcell factor for the cavity mode. We evaluate the \( F_p \) of the lasing mode to be 0.215 from Eq. (8), whereas \( F_p \) takes the value of 8.79 if evaluated using Eq. (1). In fact, Eq. (8) simplifies to Eq. (1) if the cavity lineshape is much wider than the gain medium inhomogeneous broadening lineshape, which is not the case in moderate- to high-Q semiconductor lasers. Finally, note that the Purcell factor \( F_p \) is the sum of contributions \( F_p^k \) from each cavity mode present, as is the emission probability in Eq. (4).

### 3.3 Temperature dependence of the spontaneous emission factor

In Section 2 we discussed the temperature dependence of the material gain spectra and cavity resonances. The PL spectra used in Section 3.1 and 3.2 is closely related to the gain spectra [44]; they both, for example, depend on material properties such as bandgap energy. Because of the temperature dependence of the PL spectra, as well as that of homogeneous and cavity lineshapes, the Purcell factor varies with temperature. This can perhaps best be seen by rewriting Eq. (8) with temperature made as an explicit variable.

\[
F_p^k(T) = \frac{\pi \left( \frac{c}{n_k} \right)^2 \omega_k(\omega) \Gamma_k}{\tau_{coll}} \int Z_{sp}(\omega_{21}, T) L_k(\omega - \omega_k, \tau_{coll}, T) d\omega d\omega_{21}.
\]

In Eq. (9), \( Z_{sp}(\omega_{21}, T) \) is the rate of spontaneous emission in the absence of the cavity, describing inhomogeneous broadening similar to \( D(\omega_{21}) \) in Eq. (8), but calculated according to [44]. The temperature dependence of the spontaneous emission factor, \( \beta \), follows directly from Eq. (9). We can isolate the effect of the temperature on the competition between modes by defining \( \beta_{max} \) as the ratio of the Purcell factor of the lasing mode to the sum of Purcell factors of all cavity modes,

![Figure 4](image-url) (A) The lasing mode’s electric field profile and the three spectra in the evaluation of the Purcell factor, (B) cavity lineshape, (C) homogeneous broadening lineshape and (D) PL spectrum (Reprinted from [26]).
\[ \beta_{\text{max}}(T) = \frac{E_{p}^{\text{trans}}(T)}{\sum \Delta E_{p}^{\text{r/d}}(T)}. \] (10)

To illustrate the evaluation of Eqs. (9) and (10) for a specific cavity, we identify the transparent carrier densities corresponding to the cavity modes of the resonator presented in [35], which has a total radius of 350 nm and an optimized SiO\(_2\) shield of 100 nm. If a large spontaneous emission factor is desired, this resonator is more favorable compared to the resonator of [20] because its smaller size admits fewer modes and a larger free-spectral range. The TE\(_{011}\) mode has the lowest threshold gain of the 350 nm resonator and, depending upon the temperature, its resonant wavelength ranges from ~1465 nm to ~1515 nm for positive TOC and ~1570 nm to ~1470 nm for effectively negative TOC (see Figure 3D). The transparent carrier densities are found by computing the material gain and noting the carrier density at which the gain changes sign for a particular wavelength. Over the temperature range 77 K–400 K, the transparent carrier density of the 10 nm 1.6Q/1.3Q InGaAsP QW ranges from ~2\(10^{17}\) cm\(^{-3}\) to ~3\(10^{18}\) cm\(^{-3}\), as shown in Figure 5A. The spontaneous emission factor is plotted as a function of temperature, in Figure 5B, for both positive and effectively negative TOCs.

At and above room-temperature \(\beta_{\text{max}}\) is over 0.5 because the TE\(_{011}\) mode lies close to the peak of the spontaneous emission spectrum. While the TE\(_{011}\) mode has a much higher Q than neighboring modes, the neighboring modes still overlap with the emission spectrum, preventing \(\beta_{\text{max}}\) to ever reach unity. For the case of positive TOC, the maximum of \(\beta_{\text{max}}\) occurs at 200 K when the TE\(_{011}\) mode most nearly coincides with the peak spontaneous emission. Below 200 K, however, \(\beta_{\text{max}}\) drops significantly as the emission spectrum blue-shifts and the TE\(_{011}\) mode wavelength remains nearly constant. The sharp roll-off in the MQW gain spectrum (see Figure 1A), which comes from the step-like density of states function, is responsible for the sharp decrease in \(\beta_{\text{max}}\). At low temperatures \(\lambda_{\text{TE011}}\) falls to the red side of the spontaneous emission spectrum and neighboring modes receive an increasing ratio of the spontaneous emission. The resonant wavelength of the HE\(_{211}\) mode, especially, always resides near the peak of the emission for low temperatures. Nonetheless, for the positive TOC case, the TE\(_{011}\) mode maintains sufficient overlap with the gain to remain the lasing mode for all \(T\).

For the case of effectively negative TOC, below 150 K the TE\(_{011}\) mode has negligible overlap with the gain spectrum, while the HE\(_{121}\) and HE\(_{211}\) modes have a strong overlap. In this case, it is possible that the hybrid modes may lase, depending on their threshold gains. The HE\(_{211}\) mode, in particular, has a threshold gain of \(g_{\text{n}\text{r}} < 2000\) cm\(^{-1}\) for \(T < 150\) K. Based on Figure 1A, large, but not unattainable, carrier densities would be required for the HE\(_{211}\) mode to lase. For \(T = 77\) K and effectively negative TOC, we compute \(\beta_{\text{max}} = 0.39\) for the HE\(_{121}\) mode. Due to the two-fold degeneracy of this mode, \(\beta_{\text{max}}\) cannot exceed 0.5.

Qualitatively, we see similar behavior for the cases of positive and effectively negative TOCs. The major quantitative difference results from the fact that, below 300 K, \(\lambda_{\text{TE011}}\) of the latter case exceeds that of the former. For example, at 200 K \(\lambda_{\text{TE011}}(200\) K\) = 1.483 for positive TOC, whereas \(\lambda_{\text{TE011}}(200\) K\) = 1.529 for negative TOC. As indicated by the material gain of Figure 1A, the spontaneous emission spectrum blue-shifts significantly with decreasing temperature, so the observation that the shorter wavelength has a larger \(\beta_{\text{max}}\) for a given temperature below 300 K agrees with our expectation.

At 300 K, \(\lambda_{\text{transp}}\) for effectively negative TOC is equal to \(\lambda_{\text{transp}}\) for positive TOC, leading to identical \(\beta_{\text{max}}\)’s. Above 300 K, \(\lambda_{\text{transp}}\) for negative TOC is less than \(\lambda_{\text{transp}}\) for positive

![Figure 5](image_url)  
(A) Transparent wavelength versus carrier density and (B) spontaneous emission factor versus temperature for MQW InGaAsP metal-clad nanolasers. The cases of positive and effectively negative TOCs are denoted by \(dn/dT>0\) and \(dn/dT<0\), respectively. (Reprinted from [35]).
TOC but the difference is not as significant as for low temperatures. While the peak spontaneous emission red-shifts with increasing $T$, we observe that $\beta_{\text{max}}$ for effective negative TOC is almost identical to $\beta_{\text{max}}$ for positive TOC above 300 K. This result is most likely due to the broadening of the emission with $T$. The broader emission enables more cavity modes to participate in the coupling process, leading to a less dramatic difference between the two TOC cases for $\beta_{\text{max}}$.

The major effect of the temperature in the calculation of $\beta_{\text{max}}(T)$ is to shift the gain and spontaneous emission spectra significantly. The effect of the temperature on the real parts of the cladding and semiconductor permittivities is not as dramatic. The rapid rise in $\beta_{\text{max}}$ with temperature and the subsequent broad maximum was similarly observed for VCSELS and microcavities [48, 49]. However, the critical temperature at which $\beta$ decreases sharply appears to be lower for a subwavelength cavity. The transition temperatures of [48] and [49] were approximately 240 K and 290 K, respectively. In this sense, our analysis demonstrates the rather non-intuitive result that $\beta$ of a subwavelength cavity may be more robust to temperature variation than that of a larger laser. This appears to be a direct consequence of the sparse mode density in the cavity that we have considered. Furthermore, as a consequence of the small number of total modes, $\beta_{\text{max}}$ is much higher in the metallo-dielectric lasers that we analyze, compared to those analyzed in [48] and [49].

Although we have analyzed a metallo-dielectric laser cavity supporting a photonic-type mode, the method may also be applied to lasers supporting plasmonic modes. For example, the coaxial plasmonic laser of [24], with an inner conductor radius of 100 nm and semiconductor annulus of width 100 nm (Structure A), exhibited near-unity-\(\beta\) at $T=4.5$ K. Using the results of our analysis we may make several hypotheses concerning its performance at room-temperature. Firstly, the TEM-like mode supported by this cavity was simulated to have a resonant wavelength of $\lambda_{\text{sim},A}=1.43$ $\mu$m, but lasing was observed at $\lambda_{\text{obs},A}=1.38$ $\mu$m.

The material system employed in [24] was 10 $\text{nm}$ 1.6Q/1.3Q InGaAsP MQW, so, using Figure 1A and the fact that the nearest neighbor modes are each about 200 nm from the lasing mode, we would predict the TEM-like mode to have $\beta=1$ for $T<100$ K. The red-shift of the bandgap energy with $T$, so far discussed theoretically, is corroborated by the ~160 nm shift in wavelength of the lasing mode of Structure B of [24], from $\lambda_{\text{obs},B}(4.5$ K)$=1.32$ $\mu$m to $\lambda_{\text{obs},B}(300$ K)$=1.48$ $\mu$m. The bandgap energy for 1.6Q/1.3Q InGaAsP red-shifts ~140 nm, from 0 K to 300 K. Given that the TEM-like mode has the lowest threshold gain of the two modes in Structure B, and the experimentally observed far-field radiation patterns, we can attribute the majority of the experimentally observed red-shift to the temperature dependence of the gain medium. While the analysis may explain the observations of Structure B, perhaps more significantly, it suggests that the room-temperature $\beta$ of Structure A would decrease due to the higher order mode at the simulated resonance of $\lambda_{\text{sim},A}=1.61$ $\mu$m. Furthermore if the Q of the higher order mode is large with respect to that of the TEM-like mode, then the latter may not lase at all.

Admittedly, the applicability of this analysis becomes questionable at very low temperatures, when the assumption of continuous gain and spontaneous emission spectra in the absence of a cavity may become invalid. At low enough temperatures, the intraband lifetime may become comparable to the spontaneous emission lifetime. Hence, the assumption that the carriers interact with one another on a time scale much shorter than the time between photo-emission events may no longer be valid. This means that the use of the quasi-Fermi levels becomes invalid, and we must treat the classes that contribute to the continuous spectra at high temperatures, on a class-by-class basis. It is not immediately clear how this would affect the resulting Purcell and spontaneous emission factors. For the temperatures that we have considered in this report, this is not a concern for semiconductors, but for experiments at $T=4.5$ K, more sophisticated emission spectra models may be required.

### 4 Simulation of nanolaser’s temperature performance

#### 4.1 Calculation of self-heating

In addition to the insight into the effects of temperature on nanolaser behavior described in Sections 2 and 3, the design of thermally-robust nanolasers also requires the ability to simulate the temperature performance of nanolaser designs. The internal operating temperature of a nanolaser depends not only on the ambient temperature, but also on the amount of self-heating the nanolaser experiences. For an optically-pumped nanolaser, the self-heating will depend on the power absorbed from the optical pump; most of this absorbed power is converted to heat, and only a small portion is utilized in the generation of emitted light. For an electrically-pumped nanolaser, self-heating can be calculated using the effective heat source model used in VCSELS (29. §5.3), with modification to include the heat generated from non-radiative
recombination in the active region, which is insignificant in micro- or large-scale lasers, but can play an important role in the self-heating of nanolasers.

In the modified effective heat source model, the cavity self-heating in an electrically-pumped nanolaser can be categorized into three mechanisms: 1) junction and heterojunction heating, 2) Joule heating, and 3) non-radiative recombination heating.

The first type of heating is generated at the interface of the differently doped semiconductor layers. It consists of junction heating, a term that is designated to describe the heat generated at the interfaces between the doped semiconductors and the un-doped gain region; it also consists of heterojunction heating, which accounts for the heat generated at all doped semiconductor layer interfaces. Both terms are expressed as \( I \cdot V_{\text{th}} \), where \( I \) denotes current and \( V_{\text{th}} \) is the potential difference at the \( n \)th junction. These two terms take the same form below threshold, where \( I \) is the injection current \( I_{\text{inj}} \). Above threshold, \( I \) for junction heating is clamped at the threshold current \( I_{\text{th}} \), while the heterojunction heating continues to use \( I_{\text{inj}} \).

The second type of heating is Joule heating due to the series electrical resistance in all doped semiconductor layers, and takes the form \( (I_{\text{inj}})^2 \cdot R_s \), where the resistance \( R_s \) of the \( s \)th layer is calculated using the layer’s thickness \( t_s \), cross-sectional area \( A_s \), doping concentration \( n_s \), and carrier mobility \( \mu_s \) [64], using the formula

\[
R_s = \frac{1}{n_s q \mu_s A_s t_s}
\]

where \( q \) is the carrier charge. Each doped semiconductor layer, therefore, becomes a distributed source of Joule heating.

The third type of heating is generated by non-radiative recombination inside the gain region. In nanolasers, the non-radiative recombination heating is generated by Auger recombination and surface recombination. Auger recombination is significant at high temperatures and/or high carrier densities, and surface recombination, is significant at high temperature and/or large surface-to-volume ratios, the latter being especially significant relevant to nanolasers. Therefore, the gain region becomes a distributed heat source whose power is given by the non-radiative recombination, assuming that all non-radiative energy is converted to lattice vibrations through the creation of phonons.

The above calculation of junction and heterojunction heating requires knowledge of the potential differences needed to forward bias each junction. These can be obtained using software such as SILVACO’s ATLAS, a two-dimensional electronic device simulator. Given the operating temperature and properties of the device’s constituent materials, ATLAS self-consistently solves the Poisson equation, the Schrodinger equation, and the carrier transport equations, considering Fermi-Dirac statistics, and obtains at each injection level the carrier density, the electron and hole quasi-Fermi levels, and the potential difference necessary to forward bias the junctions.

4.2 Simulation of nanolaser heat dissipation

Once the amount of nanolaser self-heating is known, the internal operating temperature of the nanolaser can be calculated using finite element software such as COMSOL’s heat transfer module. Each layer and junction can be treated as a heat source according to the effective heat source model described in Section 4.1 (or by the amount of pump absorption, for optically pumped lasers), and the transient or steady-state temperature in the laser can be obtained.

Accurate thermal analysis of a nanolaser design requires knowledge of the thermal conductivity, heat capacity, and density of each material used in the nanolaser. Since these parameters are themselves temperature-dependent, the thermal analysis should ideally include thermal feedback mechanisms to update the material parameters as the device’s temperature rises. However, experimental or experimentally-validated thermal parameters are lacking for most commonly-used nanolaser materials; temperature-dependent study of these material properties would be valuable to future nanolaser research.

Also valuable would be investigation into microscale heat transfer, as applied to nanolasers. The heat conduction models in most commercial finite element software, such as COMSOL, use macroscopic heat transfer equations, which may break down on the micro-/nano-scale. When the device dimension becomes comparable to or smaller than the mean free path of constituent materials’ heat carriers, we enter the microscale heat transfer regime [65]. Microscale conductive and radiative heat transfer in VCSELs and convective heat transfer in carbon nanotubes have been studied [65], but this has not yet been a subject of attention in the field of nanolasers.

The above analysis also does not include the effects of non-ideal Ohmic contacts, defects at material interfaces, or the effects of surface passivation on surface recombination [66]. Further refinements of nanolaser models to include these and other parameters will serve to increase the accuracy and value of nanolaser thermal simulations, as well as to suggest avenues for design improvement.
5 Nanolaser design for improved thermal performance

The ability to simulate a nanolaser’s behavior at different operating temperatures, as discussed in Sections 2 and 3, along with the ability to determine the operating temperature of a nanolaser design at a given ambient temperature and pumping level, as discussed in Section 4, are crucial to the design of nanolasers for improved thermal performance, and to the evaluation of their success.

One method of avoiding poor thermal performance is to restrict nanolasers to relatively large device sizes. Self-heating effects such as Joule heating and surface recombination become more prominent as device size shrinks, and a device that relies on heat dissipation through the semiconductor stack has a smaller conduit for heat dissipation as the radius of the stack is decreased. The analysis of CW optically-pumped nanodisk lasers operating at 45 K in [36] used this approach in a thermal simulation of device operating temperature with decreasing pedestal radius, comparing the tradeoff between increased laser heating and improved cavity Q and confinement factor (at fixed temperature). They found, for their device geometry, an optimum range of pedestal radii for which laser heating was near its minimum value and cavity Q and confinement factor were near their maximum values. Although they did not include material gain and lasing threshold in their analysis, they successfully demonstrated lasing in a fabricated device with pedestal radius in their optimum range.

Another method of improving a laser’s thermal performance is through use of materials with higher thermal conductivities, to improve the laser’s ability to dissipate heat. The authors of [36] chose InP for their CW optically-pumped nanodisk pedestal material over the AlGaAs used in previous nanodisk lasers, due to its higher thermal conductivity, although they did not quantify the amount of improvement through thermal simulation.

In addition to heat dissipation through the laser pedestal, metal-clad nanolasers have the possibility of heat dissipation via conductive heat transfer to the metal cladding. This possibility was first explored for metal-clad nanolasers in [23], which also evaluated their laser’s self-heating, operating temperature, temperature-dependent material gain and loss, and temperature-dependent optical cavity modes. Here, we detail the approach used in [23], as an example of thermally-motivated nanolaser design and performance evaluation.

The schematic of the device being analyzed is shown in Figure 6. It is a nanolaser of the electrically pumped metal-clad type. Such nanolasers have been demonstrated at 77 K under pulsed pumping in a device with a SiN₄ shield [19], at 140 K under CW pumping in a device with a SiO₂ shield [21], and most recently at room temperature under CW pumping in a device with a SiN shield [22]. The intrinsic 300 nm thick InGaAs bulk region is the active layer, and the upper (470 nm thick) and lower (450 nm thick) InP regions are the cladding layers through which carriers are injected. A highly doped n-InGaAs layer atop the upper cladding, and a p-InGaAsP layer beneath the lower cladding, form the n and p contact layers, respectively.

Two types of dielectric materials, namely SiO₂ and SiN₄, are commonly used as the shield layer in metaldielectric nanolasers and nanoscale devices in general. In electrically pumped nanolasers, for SPP mode operation, the dielectric layer is on the order of 20 nm to provide electrical insulation; SiN₄ is used because of its effectiveness as a passivation layer [19, 67]. For photonic mode operation, as discussed in Sections 2.2 and 2.3 above, a thicker shield is usually used, to minimize cavity threshold gain [39, 68]. For photonic mode operation, SiO₂ is the usual choice of shield material, since its low refractive index.

![Figure 6 A schematic of the electrically pumped nanolaser, where r_{core} is the radius of InGaAs gain layer, Δr_{upper} and Δr_{lower} are the undercut depths of the upper and lower InP pedestals, respectively. d_{shield} is the thickness of the shield layer.](image-url)
compared to SiN, yields better mode confinement [41]. However, both SiO₂ and SiN₃ largely prevent heat dissipation through the shield, due to their low thermal conductivities. Gu et al. [23] used amorphous Al₂O₃ (α-Al₂O₃) as the shield layer, which has a higher thermal conductivity than SiO₂ or SiN₃ and can be produced via thin-film deposition techniques. The thermal properties of SiO₂ [69] and Si₃N₄ [69] deposited via plasma-enhanced chemical vapor deposition, and α-Al₂O₃ deposited via atomic layer deposition (ALD) [70–72] are listed in Table 1.

In addition to choice of shield material, the authors of [23] had another, potentially thermally-relevant, parameter to choose: the diameters of the InP pedestals above and below their gain region. The ability to choose a pedestal diameter that differed from the gain diameter is thanks to a multi-step etching process [23]. For fabrication, the initial circular masks for the pillars are created by e-beam lithography with hydrogen silsesquioxane (HSQ) negative resist. Reactive ion etch (RIE) is then used to form cylindrical pillar structures as shown in Figure 7A. Next, a two-step selective wet etching process is used to etch the doped InP layers, without affecting the InGaAs gain layer, creating undercut InP pedestals. In the first step of the selective etching, the HCl:H₃PO₄ (1:4) etchant combination is used. Due to the HCl:H₃PO₄ combination’s anisotropic etching, the etch rate is slowest in the (111) plane, resulting in cone shaped regions (Figure 7B), similar to those obtained in Ref [21]. In the second step, the HCl:CH₃COOH (1:4) combination is used, whose anisotropic etch rate in the (111) plane also produces cone shaped regions, but in the opposite direction of that from the HCl:H₃PO₄ etchant combination, if used alone (Figure 7C). Therefore, applying the two chemistries sequentially with the proper ratio of etching times, vertical pedestal sidewalls can be obtained, (Figure 7D). Because the etching rate varies for different dopant types and concentrations, the upper InP layer is always narrower than the lower InP layer.

The authors of [23] chose to fabricate a device with modest undercut, as shown in Figure 7E, and, after ALD of α-Al₂O₃ in Figure 6F, to minimize anticipated self-heating in the semiconductor stack. At 77 K under CW electrical injection, the fabricated device produced a lasing mode at 1515 nm with a linewidth of 2 nm [23]. Estimating the threshold to be the current level beyond which the lasing mode linewidth stops narrowing, a signature of the onset of lasing [73], the threshold current was found to be ∼0.4 mA, while the highest experimental injection current applied was 0.5 mA. At higher temperatures, the device failed to lase.

To determine the reason for the device’s poor performance at room temperature, thermal simulations were performed at 300 K using the highest-used injection current of 0.5 mA, following the electrical-based approach described in Section 4. The rise in device temperature due to self-heating was found to be about 30 K or less, for even the most pessimistic value of α-Al₂O₃ thermal conductivity found in the literature, indicating that thermal considerations were not the performance-limiting factor for this laser.

<table>
<thead>
<tr>
<th>α-Al₂O₃</th>
<th>SiO₂</th>
<th>Si₃N₄</th>
<th>InP</th>
<th>InGaAs, AsₓP₁−ₓ, x=0.53</th>
<th>InGaAs, AsₓP₁−ₓ, x=0.773, y=0.493</th>
<th>Ag</th>
</tr>
</thead>
<tbody>
<tr>
<td>Permittivity ε</td>
<td>2.69</td>
<td>2.1</td>
<td>4.49</td>
<td>6.96</td>
<td>11.56</td>
<td>11.83</td>
</tr>
<tr>
<td>Thermal conductivity Tc (W-m⁻¹-K⁻¹)</td>
<td>1.7–20</td>
<td>1.1</td>
<td>0.7</td>
<td>68</td>
<td>16</td>
<td>11</td>
</tr>
</tbody>
</table>

Figure 7 SEM images of pedestal pillar lasers (A) after RIE, (B) after HCl:H₃PO₄ etching alone, (C) after HCl:CH₃COOH etching alone, (D) after the 2-step selective etching, (E) fabricated device after the 2-step selective etching, (F) fabricated device after α-Al₂O₃ deposition. (Reprinted from [23]).
The reason for the device’s poor performance was finally revealed through EM simulations of the as-fabricated device (using geometrical parameters measured from fabricated device SEM images) at both 77 K and 300 K, using material permittivities obtained in Section 2. For the fabricated device at both temperatures, the threshold gain \( g_{\text{th}} = \frac{2\pi n_g}{\lambda} \sqrt{\frac{Q}{\Gamma}} \), where \( n_g \) is the group refractive index and \( \Gamma \) is the mode confinement factor, was on the order of \( 10^4 \) cm\(^{-1}\). Calculation of material gain spectra, following the discussion in Section 2, reveals that to reach such high material gain values, both cryogenic temperature operation and a high carrier density are necessary.

This example highlights a case in which the combination of thermal, material, electrical, and optical analyses revealed that the most detrimental effect at high temperatures and/or under CW pumping was not the usually-suspected self-heating [74]. Nor was the laser performance limited by increased metal loss at room temperature; as described above, the electromagnetic performance of the laser was found to be the same at 77 K and 300 K, meaning that the metal loss was not the dominate loss mechanism. Instead, the most detrimental effect to the operation of the device was the angled sidewalls and the negligibly undercut lower InP pedestal, causing high radiation loss and poor mode confinement, and thus high threshold gain [23].

These insights motivated the design of an improved nanolaser for room temperature operation, in which the pedestal undercut was increased to improve the mode confinement, and thus Q-factor and threshold gain. All other laser parameters, including gain volume and sidewall angle, were kept the same. Optical simulations were performed to determine the optimal undercut depth. The improvement of optical properties of the mode with increasing undercut depth is shown in Figure 8A, which shows the threshold gain \( g_{\text{th}} \) and Q-factor of the lowest \( g_{\text{th}} \) mode, at each undercut level. The optimal undercut depth is defined to be the depth of the lower InP pedestal, \( \Delta r_{\text{lower}} \), which is the lesser of the two undercuts, that corresponds to the minimum threshold gain in Figure 8A. The optimal \( \Delta r_{\text{lower}} \) was found to be 174 nm; at \( \Delta r_{\text{lower}} \sim 200 \) nm, the original lasing mode has blue-shifted beyond the gain bandwidth window of 1300 nm to 1650 nm, leaving a higher order mode to take its place as the lowest \( g_{\text{th}} \) mode, which has a lower Q and a slightly higher \( g_{\text{th}} \) than the original mode. Figure 8B depicts the target structure’s \( \lambda_{\text{cav}} \) the electric field distribution, threshold gain \( g_{\text{th}} \) and mode confinement \( \Gamma \), for modes with \( g_{\text{th}} < 200 \) cm\(^{-1}\) and whose resonant wavelengths fall within the material electroluminescence spectrum of 1300–1650 nm. [23].

Next, the self-heating of the laser was calculated, following the approach in Section 4, to determine whether the increased undercut had adversely affected the laser’s operating temperature. The electrical and thermal simulations used the same injection current, 0.5 mA, as analyzed in the fabricated device; results are shown in Figure 9. The simulations revealed that the increased undercut had indeed increased the total amount of self-heating in the laser. However, this heat generation increase came in the form of increased Joule heating in the InP pedestals and junction heating at the lowermost heterojunction; since this heat is produced near the edges of the laser, it is more easily dissipated. On the other hand, the carrier recombination in the gain region decreased slightly. Because the new highly-undercut design had traded heating in the more thermally-isolated gain region for heating in the pedestals, the heating for this design actually decreased a fraction of a degree, compared to the as-fabricated
device, for two out of the three shield thermal conductivities simulated (Figure 9A and B). The exception was the simulation using a shield with the thermal conductivity of SiO₂ (Figure 9C). In this case, the laser was unable to dissipate heat through its shield, and instead relied on heat dissipation through the pedestal; with a smaller pedestal diameter available for heat dissipation, the laser’s internal temperature rose compared to the as-fabricated case.

Next, the optical performance of the more-undercut design was analyzed at the operating temperatures of 77 K, 300 K, and 327 K, the latter of which was the predicted steady-state temperature [assuming a modest \( T_c \) value of 10 W/(m·K) for the \( \alpha-Al_2O_3 \) shield] when operated at an ambient temperature of 300 K. According to the discussion in Section 2, only the temperature dependence of the metal cladding was considered, as the effects of temperature on the semiconductor optical parameters is expected to be negligible. The wavelength of the lowest \( g_{th} \) mode is 1376 nm with a confinement factor of 0.762, both independent of temperature as expected from the discussion in Section 2. Cavity Q-factors are 43,350, 6958, and 6487, while \( g_{th} \) values are 5 cm\(^{-1}\), 20 cm\(^{-1}\) and 31 cm\(^{-1}\), at 77 K, 300 K, and 327 K, respectively. Q and \( g_{th} \) are strongly temperature-dependent in the highly-undercut structure, because its better mode confinement reduces the amount of radiation loss and therefore makes the increase of metal loss with rising temperature more prominent.

Finally, the material gain available at the operating temperature was calculated, to check whether the device heating had significantly affected the material parameters [23]. At the 0.5 mA injection current used in the thermal and electromagnetic simulations above, SILVACO electrical simulations showed the carrier density of the designed device to be 7.072e18 cm\(^{-3}\). The gain spectra of the active medium at this injection level, at temperatures of 77 K, 300 K, and 327 K, are shown in Figure 1B. Comparing the available material gain at the carrier density of 7.072e18 cm\(^{-3}\) with the predicted threshold gain values listed in Figure 8B, even though the mode simulations do not capture the non-radiative loss that is directly related to temperature carrier density, we expect that the laser could be operated at a much lower injection level than the 0.5 mA considered, thanks to the much smaller \( g_{th} \) of the designed device compared with the fabricated one.

The above discussion shows an example of nanolaser performance analysis and design for which the interplay of thermal, optical, electrical, and material gain is essential, both for insight into the key parameters limiting device performance, and to the validation of a nanolaser design with drastically-increased performance. Additionally, the thermal effect of shield material choice was quantitatively investigated, and \( \alpha-Al_2O_3 \) shown to be a promising material to enhance heat dissipation. We expect that as the field of nanolaser design matures, thermally-based simulation will prove to be a valuable tool for the analysis and design of new nanolasers.

### 6 Discussions and conclusions

In this review, we have summarized the research efforts on temperature effects in semiconductor nanolasers, with focus on metal-clad nanolasers. The temperature effects can be explored in two paths: one path is through understanding the temperature dependence of material properties, EM mode, Purcell factor \( F_p \), as well as spontaneous emission factor \( \beta \) over a range of temperatures (77 K–400 K is considered in this review). The other path is through thermal management in device design and fabrication techniques, with the goal of achieving stable room temperature or higher temperature operation. This can be accomplished by choosing novel optically-low index and thermally-conductive materials, and considering the interplay of various design parameters including optical, electrical, thermal and material gain properties,
which builds on the understanding of temperature effects in materials and EM nano-cavities.

Because a common technique for the determination of $F_\alpha$ and $\beta$ uses curve-fitting of experimental light–light ($L - L$) curves at low temperatures in the rate equation analysis, one can more accurately analyze experimental $L - L$ curves with better understanding of the temperature dependence of $F_\alpha$ and $\beta$. The analysis can also serve as an important tool for optimizing the cavity-material system to yield maximum or minimum $\beta$, depending on the application [75]. Although we focused on analyzing an optically pumped metal-clad nanolaser with MQW gain, the analysis may be applied to an arbitrary semiconductor nanolaser geometry. It may also be applied to the construction of devices with active control of the temperature. If it were possible to tune the temperature of the device reliably, the high and low $\beta$ regions of the device might be used as a mechanism for switching between single-mode and multi-mode outputs, even before the onset of lasing.

Regarding thermal aspects of device design, we discussed techniques for simulating self-heating and heat dissipation in optically- and electrically-pumped nanolasers. We reviewed recent progress in the thermal simulation of nanolasers, including an example of a metal-dielectric laser with $\alpha$-Al$_2$O$_3$ shield. This example showed the importance of the interplay of thermal, optical, electrical, and material considerations, both for yielding insight into the key parameter limiting device performance, and to the validation of new nanolaser designs. Additionally, the analysis showed $\alpha$-Al$_2$O$_3$ to be a promising shield material, enhancing metal-clad nanolaser heat dissipation through the shield, a previously-overlooked mechanism for nanolaser heat dissipation. The integrated analysis reveals that $\alpha$-Al$_2$O$_3$’s advantages become especially evident as the pedestal undercut depth increases and/or the device size decreases.

We note that, in the thermal management studies, the thermal conductivities used for $\alpha$-Al$_2$O$_3$ represented a range of reported literature values. A 3-ω measurement can be conducted to obtain the thermal conductivity of $\alpha$-Al$_2$O$_3$ deposited under specific ALD conditions, yielding the actual experimental value used for a given laser [76]. However, the errors introduced by the two-dimensional heat spreading effect in the upper layer of the target film make this measurement non-trivial [77]. The precise determination of the heat conductivity of the deposited $\alpha$-Al$_2$O$_3$, as well as the optimization of the $\alpha$-Al$_2$O$_3$ deposition technique for increased thermal conductivity, would aid in the fabrication and analysis of future nanolasers. Recently, AlN has emerged as an effective alternative passivation material, when prepared by plasma-enhanced ALD [78]. With a higher thermal conductivity than $\alpha$-Al$_2$O$_3$, AlN can be explored as an alternative shield material in the future.

Similar to that introduced for the more mature field of VCSELs [28], a self-consistent analysis of the interconnected dynamics of nanolaser thermal behavior would incorporate the above dependencies, yielding insights into the dynamic behavior of current nanolaser designs, as well as suggesting devices with new functionalities. In combination with recent fabrication advances [23, 66], we expect the exploration of thermal effects in nanolasers to enable a new generation of robust electrically-pumped nanolasers for room-temperature operation and above.

Acknowledgments: This work was supported by the Defense Advanced Research Projects Agency (DARPA), the National Science Foundation (NSF), NSF through Center for Integrated Access Networks (CIAN) NSF ERC under grant #EEC-0812072, the Cymer Corporation, the Office of Naval Research MURI program, and the U.S. Army Research Office.

References


Graphical abstract

Qing Gu, Joseph S. T. Smalley, Janelle Shane, Olesya Bondarenko and Yeshaiahu Fainman*

**Temperature effects in metal-clad semiconductor nanolasers**

**Review article:** Schematic of an electrically pumped metal-cald nanolaser on III-V semiconductor material platform.

**Keywords:** nanoscale devices; semiconductor lasers; temperature effects.

DOI 10.1515/nanoph-2013-0058
Nanophotonics 2014; x: xxx–xxx